

RECEIVED

MAY 10 2002



TECH CENTER 1600/2900

ENTERED

1653

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/686,497

DATE: 05/02/2002 P.6

TIME: 15:58:32

Input Set : A:\10278-022001.TXT

Output Set: N:\CRF3\05022002\I686497.raw

4 <110> APPLICANT: Selden, Richard F.
 5 Miller, Allan M.
 6 Treco, Douglas A.
 8 <120> TITLE OF INVENTION: OPTIMIZED MESSENGER RNA
 10 <130> FILE REFERENCE: 10278-022001
 12 <140> CURRENT APPLICATION NUMBER: 09/686,497
 13 <141> CURRENT FILING DATE: 2000-10-11
 15 <150> PRIOR APPLICATION NUMBER: 09/407,605
 16 <151> PRIOR FILING DATE: 1999-09-28
 18 <150> PRIOR APPLICATION NUMBER: 60/130,241
 19 <151> PRIOR FILING DATE: 1999-04-20
 21 <150> PRIOR APPLICATION NUMBER: 60/102,239
 22 <151> PRIOR FILING DATE: 1998-09-29
 24 <160> NUMBER OF SEQ ID NOS: 138
 26 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 28 <210> SEQ ID NO: 1
 29 <211> LENGTH: 4376
 30 <212> TYPE: DNA
 31 <213> ORGANISM: Artificial Sequence
 33 <220> FEATURE:
 34 <221> NAME/KEY: CDS
 35 <222> LOCATION: (19)...(4353)
 37 <223> OTHER INFORMATION: synthetically generated insert
 39 <400> SEQUENCE: 1

40	tagaattcgt aggcttagc atg cag atc gag ctg agc acc tgc ttc ttc ctg	51
	Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu	
41	1 5 10	
42	15 20 25	
44	tgc ctg cgc ttc tgc ttc agc gcc acc cgc cgc tac tac ctg ggc	99
45	Cys Leu Leu Arg Phe Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly	
46	30 35 40	
48	gcc gtg gag ctg agc tgg gac tac atg cag agc gac ctg ggc gag ctg	147
49	Ala Val Glu Leu Ser Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu	
50	45 50 55	
52	ccc gtg gac gcc cgc ttc ccc ccc cgc gtg ccc aag agc ttc ccc ttc	195
53	Pro Val Asp Ala Arg Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe	
54	60 65 70 75	
56	aac acc agc gtg gtg tac aag aag acc ctg ttc gtg gag ttc acc gac	243
57	Asn Thr Ser Val Val Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp	
58	80 85 90	
60	cac ctg ttc aac atc gcc aag ccc cgc ccc tgg atg ggc ctg ctg	291
61	His Leu Phe Asn Ile Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu	
62	88 92 96	
64	ggc ccc acc atc cag gcc gag gtg tac gac acc gtg gtg atc acc ctg	339

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/686,497

DATE: 05/02/2002
TIME: 15:58:32

Input Set : A:\10278-022001.TXT
Output Set: N:\CRF3\05022002\I686497.raw

65	Gly Pro Thr Ile Gln Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu	
66	95 100 105	
68	aag aac atg gcc agc cac ccc gtg agc ctg cac gcc gtg ggc gtg agc	387
69	Lys Asn Met Ala Ser His Pro Val Ser Leu His Ala Val Gly Val Ser	
70	110 115 120	
72	tac tgg aag gcc agc gag ggc gcc gag tac gac gac cag acc agc cag	435
73	Tyr Trp Lys Ala Ser Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln	
74	125 130 135	
76	cgc gag aag gag gac gac aag gtg ttc ccc ggc ggc agc cac acc tac	483
77	Arg Glu Lys Glu Asp Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr	
78	140 145 150 155	
80	gtg tgg cag gtg ctg aag gag aac ggc ccc atg gcc agc gac ccc ctg	531
81	Val Trp Gln Val Leu Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu	
82	160 165 170	
84	tgc ctg acc tac agc tac ctg agc cac gtg gac ctg gtg aag gac ctg	579
85	Cys Leu Thr Tyr Ser Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu	
86	175 180 185	
88	aac agc ggc ctg atc ggc gcc ctg gtg tgc cgc gag ggc agc ctg	627
89	Asn Ser Gly Leu Ile Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu	
90	190 195 200	
92	gcc aag gag aag acc cag acc ctg cac aag ttc atc ctg ctg ttc gcc	675
93	Ala Lys Glu Lys Thr Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala	
94	205 210 215	
96	gtg ttc gac gag ggc aag agc tgg cac agc gag acc aag aac agc ctg	723
97	Val Phe Asp Glu Gly Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu	
98	220 225 230 235	
100	atg cag gac cgc gac gcc agc gcc cgc gcc tgg ccc aag atg cac	771
101	Met Gln Asp Arg Asp Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His	
102	240 245 250	
104	acc gtg aac ggc tac gtg aac cgc agc ctg ccc ggc ctg atc ggc tgc	819
105	Thr Val Asn Gly Tyr Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys	
106	255 260 265	
108	cac cgc aag agc gtg tac tgg cac gtg atc ggc atg ggc acc acc ccc	867
109	His Arg Lys Ser Val Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro	
110	270 275 280	
112	gag gtg cac agc atc ttc ctg gag ggc cac acc ttc ctg gtg cgc aac	915
113	Glu Val His Ser Ile Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn	
114	285 290 295	
116	cac cgc cag gcc agc ctg gag atc agc ccc atc acc ttc ctg acc gcc	963
117	His Arg Gln Ala Ser Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala	
118	300 305 310 315	
120	cag acc ctg ctg atg gac ctg ggc cag ttc ctg ctg ttc tgc cac atc	1011
121	Gln Thr Leu Leu Met Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile	
122	320 325 330	
124	agc agc cac cag cac gac ggc atg gag gcc tac gtg aag gtg gac agc	1059
125	Ser Ser His Gln His Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser	
126	335 340 345	
128	tgc ccc gag gag ccc cag ctg cgc atg aag aac aac gag gag gcc gag	1107
129	Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/686,497

DATE: 05/02/2002
TIME: 15:58:32

Input Set : A:\10278-022001.TXT
Output Set: N:\CRF3\05022002\I686497.raw

130	350	355	360	
132	gac tac gac gac gac ctg acc gac agc gag atg gac gtg gtg cgc ttc			1155
133	Asp Tyr Asp Asp Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe			
134	365	370	375	
136	gac gac gac aac agc ccc agc ttc atc cag atc cgc agc gtg gcc aag			1203
137	Asp Asp Asp Asn Ser Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys			
138	380	385	390	395
140	aag cac ccc aag acc tgg gtg cac tac atc gcc gcc gag gag gag gac			1251
141	Lys His Pro Lys Thr Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp			
142	400	405	410	
144	tgg gac tac gcc ccc ctg gtg ctg gcc ccc gac gac cgc agc tac aag			1299
145	Trp Asp Tyr Ala Pro Leu Val Ala Pro Asp Asp Arg Ser Tyr Lys			
146	415	420	425	
148	agc cag tac ctg aac aac ggc ccc cag cgc atc ggc cgc aag tac aag			1347
149	Ser Gln Tyr Leu Asn Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys			
150	430	435	440	
152	aag gtg cgc ttc atg gcc tac acc gac gag acc ttc aag acc cgc gag			1395
153	Lys Val Arg Phe Met Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu			
154	445	450	455	
156	gcc atc cag cac gag agc ggc atc ctg ggc ccc ctg ctg tac ggc gag			1443
157	Ala Ile Gln His Glu Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu			
158	460	465	470	475
160	gtg ggc gac acc ctg ctg atc atc ttc aag aac cag gcc agc cgc ccc			1491
161	Val Gly Asp Thr Leu Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro			
162	480	485	490	
164	tac aac atc tac ccc cac ggc atc acc gac gtg cgc ccc ctg tac agc			1539
165	Tyr Asn Ile Tyr Pro His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser			
166	495	500	505	
168	cgc cgc ctg ccc aag ggc gtg aag cac ctg aag gac ttc ccc atc ctg			1587
169	Arg Arg Leu Pro Lys Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu			
170	510	515	520	
172	ccc ggc gag atc ttc aag tac aag tgg acc gtg acc gtg gag gac ggc			1635
173	Pro Gly Glu Ile Phe Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly			
174	525	530	535	
176	ccc acc aag agc gac ccc cgc tgc ctg acc cgc tac tac agc agc ttc			1683
177	Pro Thr Lys Ser Asp Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe			
178	540	545	550	555
180	gtg aac atg gag cgc gac ctg gcc agc ggc ctg atc ggc ccc ctg ctg			1731
181	Val Asn Met Glu Arg Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu			
182	560	565	570	
184	atc tgc tac aag gag agc gtg gac cag cgc ggc aac cag atc atg agc			1779
185	Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser			
186	575	580	585	
188	gac aag cgc aac gtg atc ctg ttc agc gtg ttc gac gag aac cgc agc			1827
189	Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser			
190	590	595	600	
192	tgg tac ctg acc gag aac atc cag cgc ttc ctg ccc aac ccc gcc ggc			1875
193	Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly			
194	605	610	615	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/686,497

DATE: 05/02/2002
TIME: 15:58:32

Input Set : A:\10278-022001.TXT
Output Set: N:\CRF3\05022002\I686497.raw

196	gtg cag ctg gag gac ccc gag ttc cag gcc agc aac atc atg cac agc	1923
197	Val Gln Leu Glu Asp Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser	
198	620 625 630 635	
200	atc aac ggc tac gtg ttc gac agc ctg cag ctg agc gtg tgc ctg cac	1971
201	Ile Asn Gly Tyr Val Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His	
202	640 645 650	
204	gag gtg gcc tac tgg tac atc ctg agc atc ggc gcc cag acc gac ttc	2019
205	Glu Val Ala Tyr Trp Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe	
206	655 660 665	
208	ctg agc gtg ttc ttc agc ggc tac acc ttc aag cac aag atg gtg tac	2067
209	Leu Ser Val Phe Phe Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr	
210	670 675 680	
212	gag gac acc ctg acc ctg ttc ccc ttc agc ggc gag acc gtg ttc atg	2115
213	Glu Asp Thr Leu Thr Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met	
214	685 690 695	
216	agc atg gag aac ccc ggc ctg tgg atc ctg ggc cac aac agc gac	2163
217	Ser Met Glu Asn Pro Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp	
218	700 705 710 715	
220	ttc cgc aac cgc ggc atg acc gcc ctg ctg aag gtg agc agc tgc gac	2211
221	Phe Arg Asn Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp	
222	720 725 730	
224	aag aac acc ggc gac tac tac gag gac agc tac gag gac atc agc gcc	2259
225	Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala	
226	735 740 745	
228	tac ctg ctg agc aag aac gcc atc gag ccc cgc ctg gag gag atc	2307
229	Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Leu Glu Glu Ile	
230	750 755 760	
232	acc cgc acc acc ctg cag agc gac cag gag gag atc gac tac gac gac	2355
233	Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp	
234	765 770 775	
236	acc atc agc gtg gag atg aag gag gac ttc gac atc tac gac gag	2403
237	Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp Glu	
238	780 785 790 795	
240	gac gag aac cag agc ccc cgc agc ttc cag aag aag acc cgc cac tac	2451
241	Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His Tyr	
242	800 805 810	
244	ttc atc gcc gcc gtg gag cgc ctg tgg gac tac ggc atg agc agc agc	2499
245	Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser Ser	
246	815 820 825	
248	ccc cac gtg ctg cgc aac cgc gcc cag agc ggc agc gtg ccc cag ttc	2547
249	Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln Phe	
250	830 835 840	
252	aag aag gtg gtg ttc cag gag ttc acc gac ggc agc ttc acc cag ccc	2595
253	Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln Pro	
254	845 850 855	
256	ctg tac cgc ggc gag ctg aac gag cac ctg ggc ctg ctg ggc ccc tac	2643
257	Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly Pro Tyr	
258	860 865 870 875	
260	atc cgc gcc gag gtg gag gac aac atc atg gtg acc ttc cgc aac cag	2691

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/686,497

DATE: 05/02/2002
TIME: 15:58:32

Input Set : A:\10278-022001.TXT
Output Set: N:\CRF3\05022002\I686497.raw

261	Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn Gln			
262	880	885	890	
264	gcc agc cgc ccc tac agc ttc tac agc agc ctg atc agc tac gag gag	2739		
265	Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu			
266	895	900	905	
268	gac cag cgc cag ggc gcc gag ccc cgc aag aac ttc gtg aag ccc aac	2787		
269	Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn			
270	910	915	920	
272	gag acc aag acc tac ttc tgg aag gtg cag cac cac atg gcc ccc acc	2835		
273	Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr			
274	925	930	935	
276	aag gac gag ttc gac tgc aag gcc tgg gcc tac ttc agc gac gtg gac	2883		
277	Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp			
278	940	945	950	955
280	ctg gag aag gac gtg cac agc ggc ctg atc ggg ccc ctg ctg gtg tgc	2931		
281	Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys			
282	960	965	970	
284	cac acc aac acc ctg aac ccc gcc cac ggc cgc cag gtg acc gtg cag	2979		
285	His Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln			
286	975	980	985	
288	gag ttc gcc ctg ttc acc atc ttc gac gag acc aag agc tgg tac	3027		
289	Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr			
290	990	995	1000	
292	tac acc gag aac atg gag cgc aac tgc cgc gcc ccc tgc aac atc cag	3075		
293	Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln			
294	1005	1010	1015	
296	atg gag gac ccc acc ttc aag gag aac tac cgc ttc cac gcc atc aac	3123		
297	Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile Asn			
298	1020	1025	1030	1035
300	ggc tac atc atg gac acc ctg aaa ggc ctg gtg atg gcc cag gac cag	3171		
301	Gly Tyr Ile Met Asp Thr Leu Lys Gly Leu Val Met Ala Gln Asp Gln			
302	1040	1045	1050	
304	cgc atc cgc tgg tac ctg ctg agc atg ggc agc aac gag aac atc cac	3219		
305	Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His			
306	1055	1060	1065	
308	agc atc cac ttc agc ggc cac gtg ttc acc gtg cgc aag aag gag gag	3267		
309	Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu			
310	1070	1075	1080	
312	tac aag atg gcc ctg tac aac ctg tac ccc ggc gtg ttc gag acc gtg	3315		
313	Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val			
314	1085	1090	1095	
316	gag atg ctg ccc agc aag gcc ggc atc tgg cgc gtg gag tgc ctg atc	3363		
317	Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile			
318	1100	1105	1110	1115
320	ggc gag cac ctg cac gcc ggc atg agc acc ctg ttc ctg gtg tac agc	3411		
321	Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser			
322	1120	1125	1130	
324	aac aag tgc cag acc ccc ctg ggc atg gcc agc ggc cac atc cgc gac	3459		
325	Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg Asp			

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/686,497

DATE: 05/02/2002
TIME: 15:58:33

Input Set : A:\10278-022001.TXT
Output Set: N:\CRF3\05022002\I686497.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 7,8,9,10,11,12,13,14,15,16

Seq#:6; N Pos. 7,8,9,10,11,12,13,14,15,16

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/686,497

DATE: 05/02/2002
TIME: 15:58:33

Input Set : A:\10278-022001.TXT
Output Set: N:\CRF3\05022002\I686497.raw

L:414 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:417 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:2
L:1176 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1180 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:5
L:1181 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:1191 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1195 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:1196 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0